

# Asymmetric dispersal allows an upstream region to control population structure throughout a species' range

James M. Pringle<sup>a,1</sup>, April M. H. Blakeslee<sup>b,2</sup>, James E. Byers<sup>c</sup>, and Joe Roman<sup>d</sup>

<sup>a</sup>Ocean Process Analysis Laboratory, University of New Hampshire, Durham, NH 03824; <sup>b</sup>Marine Invasions Laboratory, Smithsonian Environmental Research Center, Edgewater, MD 21037; <sup>c</sup>Odum School of Ecology, University of Georgia, Athens, GA 30602; and <sup>d</sup>Gund Institute for Ecological Economics, University of Vermont, Burlington, VT 05405

Edited by David M. Karl, University of Hawaii, Honolulu, HI, and approved August 3, 2011 (received for review January 10, 2011)

**In a single well-mixed population, equally abundant neutral alleles are equally likely to persist. However, in spatially complex populations structured by an asymmetric dispersal mechanism, such as a coastal population where larvae are predominantly moved downstream by currents, the eventual frequency of neutral haplotypes will depend on their initial spatial location. In our study of the progression of two spatially separate, genetically distinct introductions of the European green crab (*Carcinus maenas*) along the coast of eastern North America, we captured this process in action. We documented the shift of the genetic cline in this species over 8 y, and here we detail how the upstream haplotypes are beginning to dominate the system. This quantification of an evolving genetic boundary in a coastal system demonstrates that novel genetic alleles or haplotypes that arise or are introduced into upstream retention zones (regions whose export of larvae is not balanced by import from elsewhere) will increase in frequency in the entire system. This phenomenon should be widespread when there is asymmetrical dispersal, in the oceans or on land, suggesting that the upstream edge of a species' range can influence genetic diversity throughout its distribution. Efforts to protect the upstream edge of an asymmetrically dispersing species' range are vital to conserving genetic diversity in the species.**

invasive species | marine genetics | phylogeography | physical oceanography

Novel genetic material can appear in a population through mutation, migration, or long-distance dispersal events which may be human-mediated. In a single well-mixed population, the evolution of the frequency of novel neutral alleles will be governed by random genetic drift, not their initial spatial distribution (1). However, spatial structure and complexity can alter this expectation. In a metapopulation linked by migration, alleles introduced into source populations are more likely to persist than those that are introduced into sinks (2–4). Many metapopulations are embedded in complex spatial systems with a preferential direction of migration (“asymmetric dispersal”). In these systems, little is known about the equilibrium frequency of novel alleles or how this frequency depends on the location where these new lineages appear.

Asymmetric dispersal is common where propagules are carried long distances by wind or water. In atmospheric, riverine, and oceanic flows, there is usually a predominant flow direction (downstream or downwind) that biases dispersal, and eddies or weather systems that slow or reverse such flow add a stochastic (and potentially upstream) component of migration. For example, many terrestrial plant species have propagules that can be dispersed by the winds (5), and a recent study of the spatial patterns of diversity in moss, liverwort, and lichen flora in the Southern Hemisphere were found to be best explained by the predominantly downwind dispersal (6). Further evidence that asymmetric dispersal can structure a species' genetic patterns in ecologically important ways can also be found in fungal rust plant

pathogens that are dispersed by winds: long-distance dispersal events can lead to the repeated introduction of genetic lineages from upwind regions into downwind ones, where they cause new disease outbreaks (7, 8).

Asymmetric dispersal is common in coastal marine systems for species with planktonic larvae. Mean currents transport most larvae downstream even as some larvae are dispersed upstream by eddies (9). In such systems, a study of the spatiotemporal evolution of the distribution of newly introduced, or locally evolved, genotypes as a function of the location of their origin will help us understand how asymmetric dispersal structures the population. It will also demonstrate which regions are most important to the genetic diversity of the overall population and how changes in these areas can impact the genetic structure of the species over the rest of its range. Theoretical work argues that novel genetic information introduced into the upstream edge of a population is likely to persist (10–13). Neutral genotypes that are introduced or arise elsewhere in the range will tend to be lost as deceased adults are replaced by migrants from farther upstream. The upstream edge of the species domain can retain new genotypes because there is little or no external immigration from farther upstream (14, 15), and so the population must be maintained by local production. This upstream region will export migrants containing these genotypes downstream, which will, in time, spread the genotype throughout the entire species' range.

Species introductions provide a unique opportunity to examine the genetic consequences of asymmetric dispersal. Because introductions in their early stages are typically not in equilibrium, they provide an opportunity to observe the evolution of gene frequencies. From an applied perspective, understanding the spread of invasive genotypes is also essential for identifying source populations and potential harbors for future invasions.

Here, we describe one such introduction of novel genotypes into a coastal marine system and document how the location and prevailing circulation patterns that govern dispersal allowed newly introduced and quite rare haplotypes to grow in frequency in the population. The European green crab, *Carcinus maenas* (hereafter *Carcinus*), is a widespread invasive species that was first observed in the eastern United States in 1817 (16) (Fig. 1). It took ~80 y for the crab to expand north of Cape Cod to the Gulf of Maine and then another 50 before it became established in

Author contributions: J.M.P., A.M.H.B., J.E.B., and J.R. designed research; J.M.P., A.M.H.B., J.E.B., and J.R. performed research; J.M.P., A.M.H.B., J.E.B., and J.R. analyzed data; and J.M.P., A.M.H.B., J.E.B., and J.R. wrote the paper.

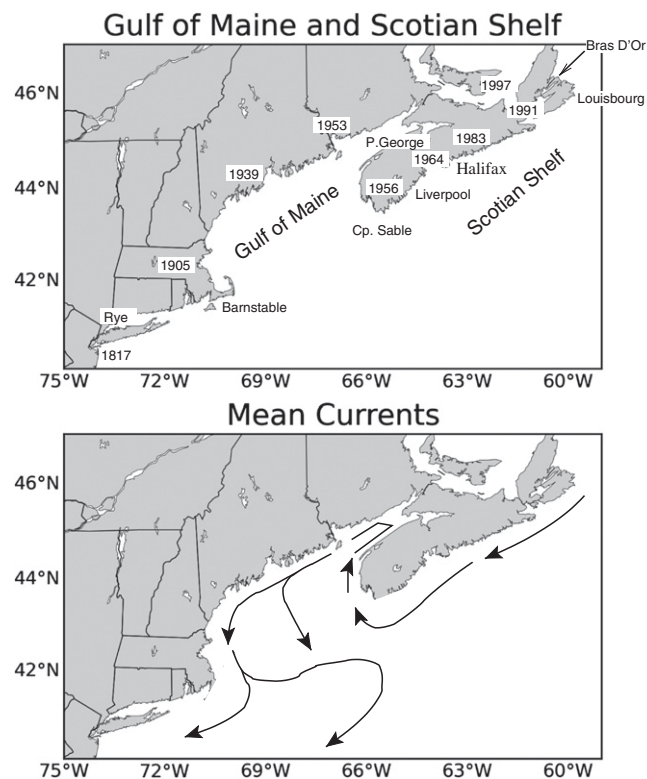
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: jpringle@unh.edu.

<sup>2</sup>Present address: Biology Department, Long Island University, Brookville, NY 11548.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100473108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100473108/-DCSupplemental).



**Fig. 1.** Map of the New England and Canadian Maritime coastline with locations mentioned in the text. (*Upper*) The dates are the years that *C. maenas* was first observed at various locations along the coast. (*Lower*) Arrows mark the mean currents in this region (data from refs. 25, 26, and 35).

Halifax Harbor. For several decades, ephemeral populations were witnessed northeast of Halifax, Nova Scotia, but they did not persist. In the 1990s, the crab displayed what appeared to be a rapid range expansion throughout the Canadian Maritimes (17). Genetic analyses revealed that these new populations of *Carcinus* were genetically distinct from downstream populations, suggesting a recent, novel introduction to northern Nova Scotia in the late 1980s or early 1990s, likely originating from northern populations in the green crab's native European range (18). The repeated introductions of *Carcinus* therefore provide an excellent natural experiment to test theories of population genetics and asymmetric dispersal.

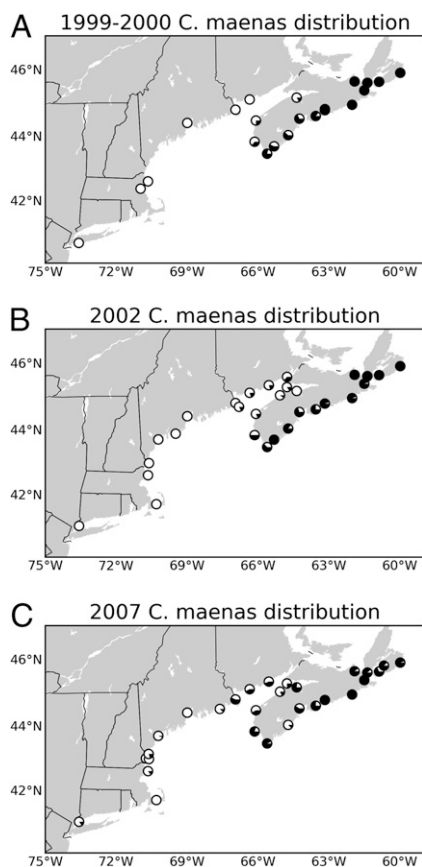
In an 8-y study of the mitochondrial haplotype distribution of *Carcinus* in the Gulf of Maine and along the Scotian Shelf of northeast North America, we examined the evolution of a cline that has formed between the population established in the 19th century and the newly introduced lineages that have been established for less than 20 y. We tracked the spatial evolution of a cline between two invasion fronts and uncovered two important patterns: (*i*) The downstream movement of the cline increases the frequency of the newly introduced upstream haplotypes and makes the composition of downstream regions more like that of the upstream edge of the species' range, and (*ii*) the downstream progress of this cline and the increase in the frequency of the upstream haplotypes are broadly consistent with the circulation in this region and the assumed neutrality of our molecular marker. In the course of this analysis, we develop a prediction for the future evolution of the cline that is robust to errors in our knowledge of larval dispersal and the demographic parameters of *Carcinus*.

## Results

To test the hypotheses that the lineages introduced at the northeast edge of *Carcinus*' range will spread downstream (southwestward) and increase their frequency in the population, we analyzed changes in haplotype frequencies across three generations of *Carcinus* along the Scotian Shelf and the Gulf of Maine. Young of the year from 2002 and 2007 were collected, sequenced, and compared with a baseline collection from 1999 to 2000 (18) (Fig. 2). The haplotypes are presented in two major classes in our analysis, one for the preexisting 19th-century lineages (southern haplotypes), and a second for the recently introduced haplotypes (northern haplotypes). Details on these groupings can be found in ref. 18 and *Materials and Methods*.

Over our 8-y timeframe, there has been a widespread increase in the recently introduced lineages. The spatial averaged frequency of the northern haplotypes between Louisbourg, Nova Scotia, and Cape Cod increased by 25% from 1999 to 2000 to 2007 ( $P < 0.01$ , Table 1). This average is not equivalent to the change of the haplotype frequency in the population because it fails to account for spatial variation in population density (*Materials and Methods*). However, the northern haplotype frequency increases from 1999 to 2007 at nearly all sampling locations except for those at the extreme northern edge of the range (Fig. 2), suggesting that the haplotype frequency in the entire population is also increasing.

Coincident with the overall increase in the frequency of the northern genotypes, the haplotype frequencies across the sampling area have become more similar to those at the upstream edge of the domain. A comparison of the level of population differentiation ( $F_{st}$ ) value from Louisbourg, at the upstream edge



**Fig. 2.** The frequency of the northern (black) and southern (white) haplotype classes for *C. maenas* for the years 1999–2000 (A), 2002 (B), and 2007 (C).

**Table 1. The along-shore distance-weighted mean allele frequency of the northern type allele calculated from Cape Cod (Barnstable, MA) to Louisbourg, Nova Scotia, along with the SE of the estimate**

Year	Frequency	Note
1999–2000	$0.36 \pm 0.014$	Mixed ages
2002	$0.39 \pm 0.012$	Young of the year
2007	$0.45 \pm 0.017$	Young of the year

The difference in frequency between 1999–2000 and 2002 is not significant ( $P < 0.10$ ), but the differences between 2007 and 2002 and 2000 are ( $P < 0.01$ ).

of the crab's range, to all downstream sampling points shows a decrease in pairwise differentiation from 1999 to 2000 to 2007 for all but one of the sampling locations (Fig. 3) (the change from 1999 to 2000 to 2002 is not significant).

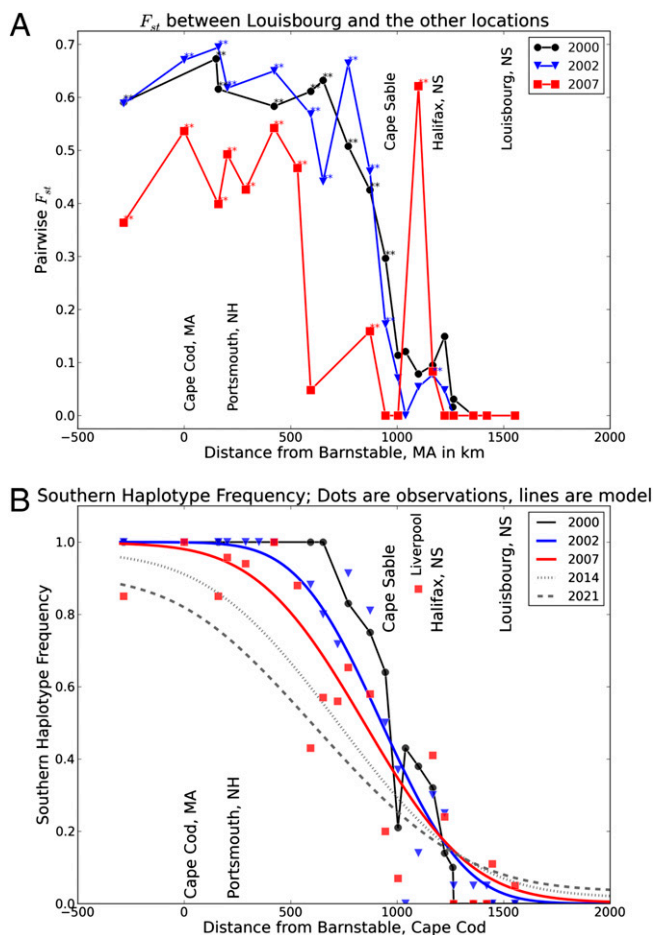
To examine whether the observed pattern can reasonably be ascribed to larval dispersal driven by coastal circulation under the

assumption of neutral genetic markers, we created a numerical model of the evolution of haplotype frequencies based on our estimates of the green crab's demographic parameters. The model uses the observed haplotype frequency in 1999–2000, 2002, and 2007 to estimate dispersal, produce optimally smoothed estimates of haplotype frequency in 2002 and 2007, and predict haplotype frequency in the years after 2007 (*Materials and Methods*). The estimates of larval dispersal are then compared in the discussion to estimates derived from the observed regional ocean circulation to determine whether the evolution of the haplotype frequency is consistent with circulation-driven asymmetric dispersal. Assuming allelic neutrality, the model estimates that the mean distance the larvae transported downstream before recruiting,  $L_{adv}$ , is  $67 \pm 17$  km, whereas the SD of the larval transport distance,  $L_{diff}$ , is  $234 \pm 19$  km.  $L_{diff}$  is greater than  $L_{adv}$ , suggesting significant upstream transport of larvae even as most move downstream. Interannual variability does not influence these estimates greatly, for we get similar estimates of  $L_{adv}$  and  $L_{diff}$  if the model is fit to only the data from 2002 or 2007 (Table 2). These estimates are robust to errors in the demographic parameters; changing the lifespan by 2 y or changing the age/fecundity relation biases the estimates of  $L_{adv}$  by  $\sim 20\%$  (Table 3; the origin of the sensitivity is described in *Materials and Methods*). We have no evidence of a selective advantage for the new lineages—all observed substitutions in these lineages are silent and in the third-codon position, and both haplotype classes are increasing in frequency at the locations where they have recently arrived. Regardless, selection of 0.1 for or against the northern haplotype relative to the southern haplotype biases estimates of  $L_{adv}$  by  $\sim 30\%$  (Table 2 and *Materials and Methods*). This dependence on demographic parameters and selection is not strong enough to change the conclusions presented below.

The observed cline between the northern and southern haplotype assemblages shifts downstream (to the southwest) and broadens with time (Figs. 2 and 3). The modeling predicts that this evolution will continue after 2007, with the frequency of the southern haplotype continuing to decrease in the south (Fig. 3) and increasing slightly in northern populations. The location of the midpoint of the cline should shift south by  $\sim 300$  km from 1999 to 2000 to 2021. This prediction is robust to errors in the demographic parameters and selection; the model compensates for any error in these parameters by biasing its estimates of larval dispersal parameters  $L_{adv}$  and  $L_{diff}$  to maintain the fit to the observations. The model's robustness is quantified by the error of the fit between the modeled and observed haplotype frequencies; "Error" in Tables 2 and 3). This error only changes slightly, and the predictions of the future cline positions through 2021 do not change, when we include moderate selection ( $s \sim \pm 0.1$ ) or when we change demographic parameters (*Materials and Methods*). In the long run, however, a favored southern haplotype would begin to grow in frequency again if, as observed, it is present at the upstream edge of the species' range.

### Discussion of the *Carcinus* System

In our 8-y investigation, we observed several noteworthy patterns that illuminate the effects of asymmetric dispersal on genetic diversity patterns in advective systems: (i) We discovered a downstream shift in the position of the center of the genetic cline between established haplotypes and those introduced within the past several decades, (ii) we observed an increase in the overall frequencies of the newly introduced haplotypes throughout the crab's range, and (iii) we witnessed a decrease in pairwise differentiation ( $F_{st}$ ) between sites across the crab's range. Along with our estimates of mean dispersal given above, these patterns are consistent with recent theories of genotype spread and clinal evolution in advective coastal ocean systems (10). The mean downstream transport of larvae,  $L_{adv}$ , drives the



**Fig. 3.** (A) Pairwise  $F_{st}$  values calculated between the northernmost sample site (Louisbourg, Nova Scotia) and all other sites for the years 1999–2000, 2002, and 2007. Higher values of pairwise  $F_{st}$  indicate greater differentiation of haplotype frequencies between Louisbourg and downstream populations ( $*P > 0.05$ ). (B) The frequency of the southern haplotype class. Markers represent observed frequencies of the southern haplotype class in 1999–2000 (black circles), 2002 (blue triangles), and 2007 (red squares). Colored lines are the model's best estimate of haplotype frequency for these years for the age classes observed; the model and data match in 1999–2000 because this data are used to initialize the model. The two gray dashed lines represent the model's predictions of haplotype frequencies in 2014 and 2021.

**Table 2. Model estimates of mean larval dispersal parameters obtained by fitting to the observed haplotype data and using the best estimates of lifespan, age/fecundity ratios, and selection**

Selection for northern assemblage	Fit to 2002 data			Fit to 2007 data			Fit to 2002 and 2007 data			Fit to 2002 and 2007 data, no outlier		
	$L_{adv}$ , km	$L_{diff}$ , km	Error	$L_{adv}$ , km	$L_{diff}$ , km	Error	$L_{adv}$ , km	$L_{diff}$ , km	Error	$L_{adv}$ , km	$L_{diff}$ , km	Error
-0.1	78	160	1.12	91	273	1.06	87	236	1.05	102	199	0.64
0	62	157	1.12	69	272	1.06	<b>67</b>	<b>234</b>	<b>1.05</b>	84	198	0.64
0.1	45	153	1.13	44	273	1.06	46	234	1.05	65	197.1	0.64

To show the sensitivity to selection, we estimate  $L_{adv}$  and  $L_{diff}$  for different values of selection, and to estimate interannual variability in the system, we present estimates from fits to data from different years. Negative values of  $s$  indicate selection against the northern haplotype assemblage, and for the southern assemblage, vice versa. Entries in boldface indicate the base parameter set presented in the article. Error is the sum of the squared difference between the observed and modeled haplotype frequencies in both 2002 and 2007. The calculations are made by fitting the model to just the 2002 data, just the 2007 data, both the 2002 and 2007 data, and both the 2002 and 2007 data with the exception of the outlier in Liverpool, Nova Scotia, in 2007. The SEs of  $L_{adv}$  and  $L_{diff}$  do not vary greatly between the various estimates, and for the base parameter set are 17 km for  $L_{adv}$  and 20 km for  $L_{diff}$ .

downstream shift in the cline between the two haplotype classes and leads to the increase in the introduced haplotypes' abundances in the population. Fewer than 20 y after the upstream establishment of the new invasion front, the dominant south-westward coastal currents (Fig. 1) have increased the frequency of the introduced haplotypes throughout the Scotian Shelf, the Gulf of Maine, and even as far as Long Island Sound, 1,800 km to the south. These population-level changes in haplotype frequencies illustrate the sensitivity of the upstream border of a species' range to anthropogenic change, whether in the form of new introductions or habitat changes that create new retention zones or allow the species range to expand upstream.

The variability in the transport of individual larvae  $L_{diff}$  increases the width of the cline (Fig. 3), leading to genetic admixture, with the southern haplotype becoming present in the northernmost sampling sites and vice versa. The magnitude of  $L_{diff}$  is greater than that of  $L_{adv}$ , indicating significant upstream larval transport, which drives the spread of the southern haplotypes into the upstream-most portion of the range; this spread is significant because the appearance of the southern haplotype class at the northern edge of the range should allow those haplotypes to persist in the population (10). Both the downstream shift in the cline and the incursion of the southern haplotype into northern populations has reduced the pairwise differentiation observed between the northern sites and the downstream locations. Left unanswered in this analysis is why stochastic larval transport had not moved the crabs with the southern haplotypes northward before the second introduction. The area just to the west of Halifax may have lacked sufficient retention zones for *Carcinus* to persist and spread to the Strait of Canso and Chedabucto Bay, currently an important retention zone. It is also possible that Allee effects before the establishment of crabs in northern regions after the second introduction prevented sporadic larval recruits from establishing populations upstream on

their own; however, our current observations cannot confirm or refute this hypothesis.

The observed downstream transport of the larvae that drives the movement of the cline is consistent with local oceanographic conditions and our understanding of coastal larval dynamics. Estimates of the regional currents (summarized in *Materials and Methods*) suggest a mean downstream transport of 280 km for a surface drifter in the core of the coastal current for the planktonic duration of the larvae; this would be the mean downstream transport of larvae that moved immediately to the core of the coastal current after spawning (~30–50 km off shore), remained at the surface, drifted in the core of that current for their entire larval duration, and then moved immediately back to shore to settle. It is implausible to expect the *Carcinus* larvae to make this exact journey with this exact timing because, at least until their later stages, the planktonic larvae cannot swim a significant distance across the shelf. Also, the larvae do not spend all of their time at the surface (19, 20). Shanks and colleagues have found the observed mean dispersal distance of larvae is usually three to five times less than what one would predict from the larval duration and the mean surface currents in the central part of the continental shelf (21, 22). Our estimate of  $L_{adv}$  is consistent with this finding.

The SD of the larval transport distance ( $L_{diff}$  in Table 2), which describes the variability of dispersal experienced by individual larvae, is higher than would be expected based on the regional circulation statistics.  $L_{diff}$  is driven by the variability in the currents experienced by larvae spawned at different times or entrained into different coastal eddies (9, 14); it could also include the effects of other larval transport mechanisms, such as shipping or fisheries. There is strong weather system-driven (23, 24) and interannual variability (25, 26) in the coastal current; its SD is comparable to the mean, suggesting that the upper limit on the stochastic component of larval transport should be the magnitude of the mean larval transport  $L_{adv}$  (14). Despite these expectations, the variability in larval transport diagnosed by the model from the evolution of the cline is ~3.5 times greater than the mean larval transport. It does not seem likely that this discrepancy is caused by larval behavior because behavior that shelters larvae from the mean flow should also shelter it from the variability in that flow. Rather, this discrepancy suggests that there is some long-distance larval dispersal that is not well captured by the statistics of the local circulation.

We cannot determine from the data whether the large stochastic long-distance dispersal (i.e., the high value of  $L_{diff}$ ) is because of the transport of larvae by unusual and extreme oceanographic conditions or by some other (perhaps anthropogenic) transport mechanism. The model is sensitive to any larval transport mechanism, and even a small number of extreme events can significantly alter the SD. For example, there is a clear

**Table 3. Model estimates of  $L_{adv}$  and  $L_{diff}$  made by models with different lifespan and age/fecundity relation**

Lifespan	Fecundity increase with age			Fecundity constant		
	$L_{adv}$ , km	$L_{diff}$ , km	Error	$L_{adv}$ , km	$L_{diff}$ , km	Error
3	52	213	1.03	42	190	1.03
5	<b>67</b>	<b>234</b>	<b>1.05</b>	55	214	1.04
7	78	248	1.07	66	231	1.06

Lifespan is varied from 3 to 7 y, and fecundity is either kept constant with age or increased with age as described in *Materials and Methods* and *SI Materials and Methods*. Entries in boldface indicate the base parameter set presented in the article. The model is fit to the 2002 and 2007 haplotype data. Error is defined in Table 2.

outlier in the data at Liverpool, Nova Scotia, where the frequency of the southern haplotype increases from <10% to >60% between 2002 and 2007 (significant at  $P < 0.05$ ). This site has characteristics that predispose it to anthropogenic transport of *Carcinus*; the 2007 sampling was in a small fishing port where both bait transport and water discharge are potential vectors. Moreover, there is strong evidence that anthropogenic transport is responsible for the spread of *Carcinus* from Nova Scotia to eastern Newfoundland, suggesting that anthropogenic transport might also be important within our study area (27). Regardless of the mechanisms contributing to stochastic dispersal, its magnitude is likely to aid in the persistence of southern haplotypes in the population.

### Asymmetric Dispersal and Population Structure

Spatially realistic models can be helpful in the study of coastal broadcast spawners, plants whose seeds are dispersed by prevailing winds, and other species influenced by asymmetric dispersal. Unlike evenly mixed populations, the frequency of neutral haplotypes that arise in the upstream portion of a species' range can increase simply because of location. These dynamics should apply to any introduced or native species with asymmetric dispersal; in the absence of a strong selection gradient (10), barrier, or other mechanism preventing migration from upstream, genetic diversity throughout the range should eventually mirror the upstream-most population. These results can aid in the management of native species: The protection of the upstream portion of a species' range will be vital in efforts to protect coastal populations and conserve genetic diversity. They also suggest that the best areas to concentrate eradication or mitigation efforts of invasive species with asymmetric dispersal would be in upstream populations because these areas are contributing larvae and genetic material to downstream populations (3, 4).

The ability of the upstream edge of a species' range to control genotype frequencies throughout its distribution suggests several counterintuitive phenomena. As climate changes, species ranges will shift. For example, along the east coast of North America, there has been a poleward shift in species ranges since the end of the last ice age, moving the range boundaries upstream against the prevailing coastal circulation (28). Range expansions are often associated with genetic bottlenecks and founder effects, such that genetic diversity is reduced in the region the population expands into (e.g., ref. 29). The reduced genetic diversity at the upstream edge will then reduce the genetic diversity downstream in the rest of the range, suggesting that upstream range expansion in asymmetrically dispersing species might potentially result in a reduction in genetic diversity throughout the population. This is an important consideration not only for invasive species, such as *Carcinus*, but also for native species with asymmetric dispersal—such declines in genetic diversity could result in adverse effects in native populations, which may already be strained by other anthropogenically induced disturbances such as habitat degradation and destruction, pollution, overharvesting, and, of course, invasive species.

Because changes in allele frequency in the upstream portion of an asymmetrically dispersing species' range will be reflected elsewhere in the range, alleles that are favored at the upstream edge of an asymmetrically dispersing species will be likely to grow in frequency in the entire population. In systems where climate change is expected to move species range limits upstream (i.e., where the upstream direction is poleward), this expansion might tend to push the evolution of the species everywhere in its range toward traits that are favored in the region into which it is expanding.

The observations presented above have shown how asymmetric dispersal has increased the impact of a secondary introduction of an invasive species on its population genetics. We have discussed

how it allow changes in a species range to modify its genetics throughout its range. Ultimately, our observations suggest that asymmetric migration allows any processes that affect the population genetics of a species at the upstream edge of its range to alter the population genetics of the species everywhere.

### Materials and Methods

**Genetic Data.** J.R. collected adult and juvenile *Carcinus* in 1999 and 2000 from 23 sites throughout the species' range in the Northwest Atlantic (18). In late summer 2002 and 2007, young-of-the-year juvenile crabs were collected approximately every 50 km along the northern Mid-Atlantic Bight, the Gulf of Maine, and the Scotian Shelf. Thirty locations each were sampled between Rye, NY, and Louisbourg, Nova Scotia, in both 2002 and 2007 (see *SI Materials and Methods* for details of the collection and sequencing of the mitochondrial CO1 gene).

**Data Analysis.** Haplotypes were grouped into two classes as described in ref. 18. The southern haplotype class, shown in white in Fig. 2, includes haplotypes that were part of the original founding population and is found in the southern part of the range and in museum samples collected in southern Nova Scotia before 1976. The newly introduced northern haplotypes, shown in black in Fig. 2, were introduced on the northern edge of the range in the late 20th century. The frequency of the haplotype classes is given for each year and location in Fig. 2. Summary statistics for genetic variability are available in *Supporting Information*, and details of the mapping of this data to alongshore distance and the Monte Carlo error estimation of allele frequency are given in *SI Materials and Methods*.

The fraction of the southern haplotype is shown as a function of alongshore distance in Fig. 3. To examine temporal change in the genetic cline recorded in 1999–2000, 2002, and 2007, we calculated pairwise differences with Arlequin version 2.0 (30) to assess change in population structure at all locations. Pairwise  $F_{st}$  values were calculated with the complete (non-grouped) haplotype data. Confidence limits in  $F_{st}$  were obtained by the default error estimation in Arlequin.

**Population Modeling.** The numerical model is based on the work of Pringle and Wares (10, 11) and simulates the downstream motion of the cline between the two haplotype assemblages. This spatially explicit model allowed for an age-structured population with age-dependent fecundity and density dependence caused by habitat limitation. It can include selection in favor of either haplotype assemblage. The model parameterized the larval dispersal in terms of the mean distance the larvae that successfully recruited are dispersed downstream ( $L_{adv}$ ) and the SD of their dispersal distance ( $L_{diff}$ ). Because of the great uncertainty in the alongshore variability of the population density, especially the subtidal population (31), we assume that the population density is uniform. In the model runs, the dispersal kernel is assumed to be Gaussian (9); qualitatively identical results were obtained with a Laplacian dispersal kernel. The model is initialized with the distribution of the allele classes observed in 1999–2000; these observations mixed individuals from all year classes, and so the model is initialized with the same allele frequency for all ages and the population at its carrying capacity.

The details of the demographic and life-history parameters used in the model are given in *SI Materials and Methods*. The demographic parameter with the largest impact on the model's prediction of the cline's evolution is the maximum age of the adults. The origin of this error is described in *SI Materials and Methods*, and a doubling of lifespan roughly halves the speed of downstream advance of a cline (cf. refs. 32 and 33).

Any selection for one of the haplotype assemblages would also change the speed the cline moves. As discussed in *SI Materials and Methods*, if the northern haplotype assemblage (or those genes associated with it) is favored everywhere in the domain, it will cause the cline to move southwestward faster, and if they are selected against, it will tend to cause the cline to move downstream more slowly, or even upstream.

The model is then fit to the observed haplotype frequencies to estimate the larval dispersal parameters that produce the best match between the modeled and observed haplotype frequencies. Powell's method is used to find the values of mean larval transport  $L_{adv}$  and the stochastic component of larval transport  $L_{diff}$  that minimizes the sum of the square of the difference between the modeled and observed haplotype distributions in 2002 and 2007 (34). Because the 2002 and 2007 sampling focused on the young of the year, we compare modeled young-of-the-year haplotype distributions to the data. Results are shown in Tables 2 and 3. Random errors in  $L_{adv}$  and  $L_{diff}$  attributable to sampling statistics are estimated with a Monte Carlo

resampling of each location 1,000 times from a binomial distribution with the same expected value and number of samples as the observations.

We cannot precisely quantify the errors in our estimate of  $L_{adv}$  and  $L_{diff}$  because of errors in our understanding of demographics or selection, for we do not have an a priori estimate of the uncertainty in our knowledge of these parameters. However, because the model finds the  $L_{adv}$  and  $L_{diff}$  that provide the best fit to the observed haplotype data, any error in our understanding of demographic parameters or selection will lead to errors in our estimate of  $L_{adv}$  and  $L_{diff}$  while keeping the evolution of the haplotype frequency consistent with the observations. Errors in parameters that tend to slow the downstream motion of the cline, e.g., a longer lifespan or selection for the southern haplotype assemblage, would lead to an increase in the model's estimate of  $L_{adv}$  to compensate, whereas the converse would happen if we underestimated the lifespan of *Carcinus* or if there were selection for the northern haplotype assemblage. The sensitivity of the model's estimates of  $L_{adv}$  and  $L_{diff}$  to these errors is shown in Tables 2 and 3. However, even if we err in our estimate of the demographic parameters of *Carcinus*, the model still produces estimates of the past and future distribution of the haplotypes that do not change. This lack of change is seen in the sum of the squared difference between the modeled and observed haplotype distributions, the error in Tables 2 and 3. This error does not increase significantly, and the predictions of the future cline positions do not change, when we include selection (Table 2) or when we change demographic parameters (Table 3).

**Expected Larval Dispersal Parameters from the Regional Ocean Circulation.** The expected mean and SD of the *Carcinus* larval transport distance can be roughly estimated for the Scotian Shelf and Gulf of Maine under rather idealized models of larval behavior. Although these estimates are very rough, because of the uncertainty in our knowledge of the larval behavior, they form a comparison point to the estimates found by fitting the haplotype evolution model to the observed haplotype data. The details of the regional oceanography and larval planktonic durations used to estimate these distances are given in *SI Materials and Methods*.

The mean downstream transport of larvae  $L_{adv}$  would be 280 km if the larvae moved immediately to the core of the coastal current after spawning ( $\sim 30$ – $50$  km off shore, over the 60- to 100-m isobaths), drifted in the core of the coastal current for their entire larval duration, then moved immediately back to shore on settle. However, it is quite implausible to expect the planktonic larvae to be able to move immediately to the core of a current tens of kilometers off shore and then to return promptly when they are ready to recruit. Thus, we must treat 280 km as an upper bound on their transport distance and expect the actual larval transport to be smaller, perhaps quite significantly so. Comparisons of mean larval transport estimated from mean currents to indirect estimates of larval transport suggest that the estimates from midshelf currents tend to overestimate the mean transport distance by a factor of 3–5 (22).

The SD of the larval dispersal distance  $L_{diff}$  of the larvae released by an adult is driven by two distinct phenomena: (i) the dispersal of larvae released in a single dispersal event by eddies and fluctuations in the mean currents (9) and (ii) the varying dispersal distance of larvae released by a single adult during different seasons or different years (9, 14). The larger of these two will dominate the total  $L_{diff}$ . Details of the estimate of  $L_{diff}$  from the variable circulation in the Gulf of Maine and Scotian Shelf are given in *SI Materials and Methods*; they show that the estimate of  $L_{diff}$  is dominated by interannual and intraseasonal variability in  $L_{adv}$  and should be of the same order as  $L_{adv}$ .

**ACKNOWLEDGMENTS.** The original data collection and preliminary analyses were performed by J.R. in the laboratory of S. Palumbi, who was essential in catalyzing this research. We thank Carrie Keogh for help with data collection, Caitlin Hamer for field and laboratory assistance, and Michael Lesser for use of equipment and facilities. A.M.H.B. thanks Greg Ruiz, Whitman Miller, the Invasions laboratory, and the Plant Ecology laboratory at the Smithsonian Environmental Research Center for support and use of facilities. The 2007 data collection was supported by the Sloan Foundation's Census of Marine Life. J.M.P. and J.E.B. were supported by National Science Foundation Grant OCE0961344.

- Hartl DL (1981) *A Primer of Population Genetics* (Sinauer, Sunderland, MA).
- Kawecki TJ, Holt RD (2002) Evolutionary consequences of asymmetric dispersal rates. *Am Nat* 160:333–347.
- Vuilleumier S, Possingham HP (2006) Does colonization asymmetry matter in metapopulations? *Proc Biol Sci* 273:1637–1642.
- Lundy IJ, Possingham HP (1998) Fixation probability of an allele in a subdivided population with asymmetric migration. *Genet Res* 71:237–245.
- Tackenberg O (2003) Modeling long-distance dispersal of plant diaspores by wind. *Ecol Monogr* 73:173–189.
- Muñoz J, Felicísimo AM, Cabezas F, Burgaz AR, Martínez I (2004) Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* 304:1144–1147.
- Nagarajan S, Singh DV (1990) Long-distance dispersion of rust pathogens. *Annu Rev Phytopathol* 28:139–153.
- Brown JKM, Hovmöller MS (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* 297:537–541.
- Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion. *Mar Ecol Prog Ser* 260:83–96.
- Pringle JM, Wares JP (2007) Going against the flow: Maintenance of alongshore variation in allele frequency in a coastal ocean. *Mar Ecol Prog Ser* 335:69–84.
- Wares JP, Pringle JM (2008) Drift by drift: Effective population size is limited by advection. *BMC Evol Biol* 8:235.
- Nagylaki T (1978) Clines with asymmetric migration. *Genetics* 88:813–827.
- Morrissey MB, de Kerckhove DT (2009) The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. *Am Nat* 174:875–889.
- Byers JE, Pringle JM (2006) Going against the flow: Retention, range limits and invasions in advective environments. *Mar Ecol Prog Ser* 313:27–41.
- Largier J (2004) The importance of retention zones in the dispersal of larvae. *Proc American Fish Society Symp* 42:105–122.
- Say T (1817) An account of the Crustacea of the United States. *J Acad Nat Sci Phila* 1: 56–57.
- Audet D, et al. (2003) Geographical expansion of a nonindigenous crab, *Carcinus maenas* (L.), along the Nova Scotian shore into the southeastern Gulf of St. Lawrence, Canada. *J Shellfish Res* 22:255–262.
- Roman J (2006) Diluting the founder effect: Cryptic invasions expand a marine invader's range. *Proc Biol Sci* 273:2453–2459.
- Queiroga H (1996) Distribution and drift of the crab *Carcinus maenas* (L.) (Decapoda, Portunidae) larvae over the continental shelf off northern Portugal in April 1991. *J Plankton Res* 18:1981–2000.
- Peliz A, et al. (2007) A study of crab larvae dispersal on the Western Iberian Shelf: Physical processes. *J Mar Syst* 68:215–236.
- Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. *Ecol Appl* 13:159–169.
- Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biol Bull* 216: 373–385.
- Beardsley RC, Chapman DC, Brink KH, Ramp SR, Schlitz R (1985) The Nantucket Shoals Flux Experiment (NSFE79). Part I: A basic description of the current and temperature variability. *J Phys Oceanogr* 15:713–748.
- McMullin G (2003) Longshore momentum balance for a mid-shelf region of the Gulf of Maine using the Gulf of Maine Ocean Observing System. MA thesis (University of New Hampshire, Durham, NH).
- Pettigrew NR, et al. (2005) The kinematic and hydrographic structure of the Gulf of Maine Coastal Current. *Deep Sea Res Part II Top Stud Oceanogr* 52:2369–2391.
- Pringle JM (2006) Sources of variability in Gulf of Maine circulation, and the observations needed to model it. *Deep Sea Res Part II Top Stud Oceanogr* 53:2457–2476.
- Blakeslee AMH, et al. (2010) A hitchhiker's guide to the Maritimes: Ocean transport facilitates long-distance dispersal of a marine crab to Newfoundland. *Divers Distrib* 16:879–891.
- Loder JW, Petrie B, Gawarkiewicz G (1998) The coastal ocean off northeastern North America: A large-scale view. *The Sea* 11:105–133.
- Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annu Rev Ecol Syst* 40:481–501.
- Schneider S, Roessli D, Excoffier L (2000) *Arlequin: A Software for Population Genetics Data Analysis* (Genetics and Biometry Lab, University of Geneva, Geneva), Version 2, 2496–2497.
- Klassen G, Locke A (2007) A biological synopsis of the European green crab, *Carcinus maenas*. *Can Manuscr Rep Fish Aquat Sci* 2818:1–75.
- Lutscher F (2007) A short note on short dispersal events. *Bull Math Biol* 69:1615–1630.
- Pringle JM, Lutscher F, Glick E (2009) Going against the flow: Effects of non-Gaussian dispersal kernels and reproduction over multiple generations. *Mar Ecol Prog Ser* 377: 13–17.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) *Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ Press, Cambridge, UK), 3rd Ed.
- Hannah CG, Shore JA, Loder JW, Naimie CE (2001) Seasonal circulation on the Western and Central Scotian Shelf. *J Phys Oceanogr* 31:591–615.